The highlight of the steps took to bring mod03's Project to completion.

1-Importing the libraries

We have imported the necessary libraries: sqlite3, numpy, pandas, scipy.stats as scs, from statsmodels.stats.power we imported tt_ind_solve_power, matplotlib.pyplot as plt and seaborn as sns

2-connected to the database

We used this code to connect to the database:

conn = sqlite3.connect('Northwind_small.sqlite')

cursor = conn.cursor()

3-Listed all the tables in the database with this code:

tables = cursor.execute("SELECT name FROM sqlite_master WHERE type='table';").fetchall()

tables = [i[0] for i in tables]

tables

4-We have aim for the table that will give us must of the information important for our project.

OrderDetail has the most relevant information.

5-Methodology

We have used Weltch test, Cohen'd, and Anova for our hypothesis test.

For each hypothesis we have specified the null and alternate hypothesis: here are the four pair of hypothesizes that we used:

H0: the discount does not have any influence on the quantity of order

Ha: the quantity ordered increase when discount is applied.

H0: There is no significant difference between the discounts given by employees from the USA versus the employees from UK.

Ha: There is a significant difference between the discounts given by employees from the USA versus the employees from UK.

H0: there is no statistically significant difference between the quantity of product sold in the USA versus the UK.

Ha: there is a statistically significant difference between the quantity of product sold in the USA versus the UK.

H0: there is no statistically significant difference between the processing time per order in the USA versus the UK.

Ha: there is a statistically significant difference between the processing time per order in the USA versus the UK.

For each hypothesis we were able to either confirm or infirm the null hypothesis using one of our methodology.

More research can further be done such as:

-Performance evaluation of the shipping companies involved.

-Diversity of the customer based on their location